

PCT



WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro



INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6 : G10L 5/06, 3/00		A1	(11) Internationale Veröffentlichungsnummer: WO 99/27524
			(43) Internationales Veröffentlichungsdatum: 3. Juni 1999 (03.06.99)
(21) Internationales Aktenzeichen: PCT/DE98/03366 (22) Internationales Anmeldedatum: 16. November 1998 (16.11.98) (30) Prioritätsdaten: 197 51 739.0 21. November 1997 (21.11.97) DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): HAUENSTEIN, Alfred [DE/DE]; Carl-Orff-Bogen 89, D-80939 München (DE). (74) Gemeinsamer Vertreter: SIEMENS AKTIENGESELLSCHAFT; Postfach 22 16 34, D-80506 München (DE).		(81) Bestimmungsstaaten: US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Veröffentlicht Mit internationalem Recherchenbericht. Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.	

(54) Title: METHOD AND DEVICE FOR VOICE RECOGNITION

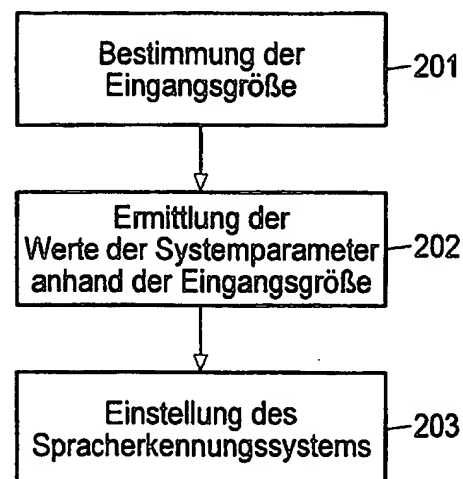
(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR SPRACHERKENNUNG

(57) Abstract

The invention relates to a method and device which permits an adjustment of precision of a voice recognition system during a voice recognition. To this end, system parameters of the voice recognition system are established using an input quantity by means of a representation specification and the voice recognition system is accordingly adjusted to the values which are determined in this manner. An adaptation of a voice recognition system can optionally result during operation in order to guarantee an application dependent matching in a range between maximum quality during voice recognition and the highest possible speed during execution of the voice recognition. The computing element can also be automatically adjusted by determining a performance index of the computing element beforehand using a program provided for the adjustment. Said performance index represents a measure for the input quantity and, as a result, guarantees an adequate operation of the voice recognition system on this computing element.

(57) Zusammenfassung

Es werden ein Verfahren und eine Vorrichtung angegeben, die bei einer Spracherkennung eine Einstellung einer Genauigkeit eines Spracherkennungssystems ermöglichen. Dazu werden anhand einer Eingangsgröße über eine Abbildungsvorschrift Systemparameter des Spracherkennungssystems ermittelt und das Spracherkennungssystem entsprechend der auf diese Art bestimmten Werte eingestellt. Optional kann eine Adaption eines Spracherkennungssystems während des Betriebs erfolgen, um eine anwendungsabhängige Anpassung in einem Bereich zwischen maximaler Qualität bei der Spracherkennung und möglichst hoher Geschwindigkeit bei der Durchführung der Spracherkennung zu gewährleisten. Auch kann automatisch der Rechner eingestellt werden, indem zuvor anhand eines dafür vorgesehenen Programmes ein Leistungsindex des Rechners bestimmt wurde, der ein Mass für die Eingangsgröße darstellt und somit einen adäquaten Betrieb des Spracherkennungssystems auf diesem Rechner gewährleistet.



201...DETERMINATION OF THE INPUT QUANTITY  
202...ESTABLISHING THE VALUES OF THE SYSTEM  
PARAMETER WITH THE AVAILABILITY OF THE  
INPUT QUANTITY  
203...ADJUSTMENT OF THE VOICE RECOGNITION  
SYSTEM

# **LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland		Republik Mazedonien	TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von
CA	Kanada	IT	Italien	MX	Mexiko		Amerika
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neuseeland	ZW	Zimbabwe
CM	Kamerun		Korea	PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

Beschreibung**Verfahren und Vorrichtung zur Spracherkennung**

- 5 Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Spracherkennung.

Ein Spracherkennungssystem ist aus [1] bekannt. Dort finden sich auch eine grundlegende Einführung der an dem  
10 Spracherkennungssystem beteiligten Komponenten sowie wichtiger, bei der Spracherkennung üblicher Techniken.

Bei einem bekannten Spracherkennungssystem ist eine Genauigkeit, also ein Maß für eine Qualität der Erkennung,  
15 vorgegeben. Der Benutzer muß nun mit diesem System auskommen, auch wenn für seine Anwendung eine verminderte Genauigkeit ausreichen, dafür er aber eine höhere Bearbeitungsgeschwindigkeit erzielen würde.

20 Das Prinzip des Pruning eines Suchraums ist aus [2] (siehe Kapitel 3.3.3, Seite 40) bekannt. Dabei handelt es sich um ein "Beschneiden" des Suchraums, also eine Methode zur Reduzierung einer Anzahl von Suchpfaden des Suchraums, wobei diejenigen Suchpfade abgeschnitten werden, die wenig  
25 aussichtsreich sind. Zuerst wird dazu ein Suchpfad mit minimalen Kosten (optimaler Suchpfad) ermittelt. Daraufhin werden alle Suchpfade (Äste des Suchbaums) weggeschnitten, deren Kosten oberhalb des Minimums zuzüglich einer addierten vorgegebenen Bewertungsgröße, die als Pruning-Schwelle  
30 bezeichnet wird, liegen. Für eine detaillierte Erklärung des Prunings: [2], Seite 40ff., insbesondere Bild 16 auf Seite 41.

Bei Verwendung der Pruning-Schwelle ist nicht bekannt, wieviel Suchpfade in dem Suchbaum übrig bleiben. Will man die  
35 Anzahl dieser übrigbleibenden Suchpfade auf einem vorgegebenen Niveau halten, wird die Pruning-Schwelle dynamisch angepaßt.

Ein Histogramm-Pruning ist aus [3] oder [4] bekannt. Hier werden eine vorgegebene Anzahl "bester" Suchpfade, also Suchpfade mit einer hohen Auftrittswahrscheinlichkeit, verwendet, indem Häufigkeiten der Suchpfade in Form eines Histogramms bewertet werden. Die Pruning-Schwelle wird dynamisch verändert.

Eine akustische Vorausschau im Suchbaum (Fachwort: Fast-Look-Ahead) ist aus [5] oder [6] bekannt.

Die bei der akustischen Vorausschau (auch schnelle Vorauswahl) verfolgte Idee beruht auf der Eigenschaft einer Sprache, daß sich alle Wörter aus einem beschränkten Inventar von Unterworteinheiten (z.B. Phonemen, Halbsilben) zusammensetzen. Für diese Unterworteinheiten wird nun "im Voraus" eine akustische Bewertung durchgeführt. Es werden nur diejenigen Kombination von Unterworteinheiten weiterverfolgt, deren akustische Bewertungen jeweils unterhalb einer vorgegebenen Schwelle liegen. Ein Gewinn im Bewertungsaufwand besteht darin, daß für eine geringe Anzahl von Unterworteinheiten ein Maß für die Übereinstimmung eines zu erkennenden Sprachsignals mit einer Zielgröße im voraus berechnet und als Grundlage für eine Entscheidung herangezogen wird, ob ein Großteil des Suchbaums nicht weiter berücksichtigt werden soll. Anschaulich gesprochen bedeutet dies, daß mehr Suchpfade im Suchbaum eingespart werden, als durch die Vorausberechnung hinzukommen. Ein derartiger Gewinn wird umso größer, je höher das Verhältnis von neuen Wortanfängen zu einer Anzahl von Unterworteinheiten wird. Dieses Verhältnis steigt mit der Anzahl der zu erkennenden Unterworteinheiten bzw. Wörter (Lexikongröße).

Ein Vorteil des Verfahrens der akustischen Vorausschau besteht in der Regularität der Algorithmen zur Berechnung der entsprechenden Maße. Da keine Verzweigungen durch Wortenden, Syntaxknoten, etc. im Suchraum auftreten, ist das Schema der Berechnung der Maße regulär. Gerade deshalb bietet sich ein

solches Verfahren auch für eine Implementierung in Hardware an.

Die Vorausberechnung der Maße (Fachwort: Fast-Match-Scores) wird dadurch möglich, daß die eigentliche Suche um eine feste Anzahl von Zeitfenstern hinter den aktuellen extrahierten Maßen des Sprachsignals hinterhereilt. Mit den aktuellen Maßen wird die Vorausberechnung der Maße weiterer Unterworteinheiten durchgeführt (siehe [5], Seite 65, Bild 33).

Auch in Sprachmodellen ist die Durchführung einer derartigen Vorausschau bekannt (siehe [6]).

Das Prinzip der Vorausschau im Sprachmodell (Fachwort: Language-Model-Look-Ahead) ist die Berücksichtigung der im Sprachmodell vorhandenen Wahrscheinlichkeiten in dem Suchprozeß so früh wie möglich, auch in dem assoziierten Pruning. Dies wird erreicht durch eine Faktorisierung der Wahrscheinlichkeiten im Sprachmodell. Eine detaillierte Beschreibung mit einer formalen Notation ist in [6] enthalten.

Schließlich ist z.B. aus [7] eine Schwelle zur Auswahl berechnender Distanzparameter bekannt. Derartige Auswahlverfahren sind generell mehrstufig angelegt. Zuerst wird eine grobe Berechnung mit einem Teil der Distanzen durchgeführt. Im nächsten Schritt werden dann diejenigen Distanzen bestimmt, die bezüglich eines Abstandsmaßes nahe an der besten Distanz des ersten Berechnungsschrittes liegen. Dieses Abstandsmaß kann über eine Schwelle variiert werden, wodurch der Berechnungsaufwand für die Bestimmung der Distanzparameter variiert wird.

Die Aufgabe der Erfindung besteht darin, ein Verfahren und eine Vorrichtung zur Einstellung der Genauigkeit des Spracherkennungssystems zu schaffen.

Diese Aufgabe wird gemäß den Merkmalen der unabhängigen Patentansprüche gelöst.

Es wird ein Verfahren zur Einstellung einer Genauigkeit eines  
5 Spracherkennungssystems angegeben, bei dem durch eine  
vorgebbare Eingangsgröße die Genauigkeit bestimmt wird.  
Anhand dieser Eingangsgröße werden Werte für Systemparameter  
des Spracherkennungssystems, vorzugsweise mittels eines  
Rechners, ermittelt. Anhand dieser Werte wird das  
10 Spracherkennungssystem eingestellt. Dies geschieht  
vorzugsweise automatisch durch den Rechner.

Somit ist es ein Vorteil der Erfindung, die Genauigkeit des  
Spracherkennungssystems adaptierbar und zugleich für den  
15 Laien einstellbar zu machen. Je nach Anwendung bzw. je nach  
Rechenleistung, die für das Spracherkennungssystem auf dem  
Rechner bereitsteht, können unterschiedliche Anforderungen an  
die Qualität des Spracherkennungssystems einfach durch  
Adaption der Eingangsgröße vorgenommen werden.

20 Eine Weiterbildung besteht darin, daß gemäß einer  
Abbildungsvorschrift aus der Eingangsgröße die Werte für die  
Systemparameter des Spracherkennungssystems ermittelt werden.  
Dabei kann diese Abbildungsvorschrift anhand einer Tabelle  
25 umgesetzt sein.

Es ist also möglich, durch Einstellung der Eingangsgröße  
automatisch die damit verknüpften Werte der Systemparameter  
des Spracherkennungssystems zu ermitteln und automatisch dem  
30 Spracherkennungssystem zugänglich zu machen. Das Ablegen der  
Werte in einer Tabelle hat den Vorteil, daß eine individuelle  
Anpassung verschiedener Werte der Eingangsgröße zu jeweils  
verschiedenen Werten der Systemparameter durchgeführt werden  
kann.

35 Eine andere Weiterbildung besteht darin, die Einstellung  
während des Betriebs des Spracherkennungssystems

durchzuführen. Dabei ergibt sich vorteilhaft, daß die Anpassung des Spracherkennungssystems individuell nach den jeweiligen Anforderungen während des Betriebs des Spracherkennungssystems angepaßt werden kann.

5

Eine zusätzliche Weiterbildung der Erfindung besteht darin, daß mindestens einer der folgenden Systemparameter anhand der Eingangsgröße bestimmt wird:

- a) Pruning-Schwelle;
- 10 b) Histogramm-Pruning;
- c) akustische Vorausschau;
- d) Vorausschau im Sprachmodell;
- e) Schwelle zur Auswahl zu berechnender Distanzparameter.

- 15 Zur Bedeutung und Funktion dieser Systemparameter wird auf die Einleitung verwiesen.

Die angeführten Systemparameter stellen eine Auswahl von Möglichkeiten dar. Es sind andere Systemparameter  
20 vorstellbar, die, abhängig von dem jeweiligen Spracherkennungssystem, von den oben genannten verschieden sein können.

Auch ist es eine Weiterbildung der Erfindung, daß die  
25 Systemparameter hinsichtlich ihres Einflusses auf eine Zielgröße hin gewichtet werden. Dabei kann die Zielgröße beispielsweise eine Genauigkeit des Spracherkennungssystems oder eine Geschwindigkeit des Spracherkennungssystems (also die Geschwindigkeit zur Durchführung des  
30 Spracherkennungsprozesses) sein. Im Hinblick auf die jeweilige Zielgröße können die Systemparameter anteilig gleich oder entsprechend einer vorgegebenen Gewichtungstabelle unterschiedlich gewichtet werden.

- 35 Im Rahmen einer anderen Weiterbildung wird die Eingangsgröße anhand eines Einstellelements bestimmt.

Vorzugsweise weist das Einstellelement einen eindimensionalen Freiheitsgrad mit zwei Begrenzungen auf, wobei die erste Begrenzung als maximale Genauigkeit des Spracherkennungssystems und die zweite Begrenzung als maximale Geschwindigkeit des Spracherkennungssystems umgesetzt werden.

Auch ist es eine Weiterbildung des Verfahrens, daß das Einstellelement auf einem Rechner als ein Schieberegler dargestellt wird und anhand einer Tastatur, eines Touch-Pads oder einer Maus bedient wird.

Auch kann das Einstellelement ein Drehregler, ein Schieberegler oder Potentiometer sein.

In einer zusätzlichen Weiterbildung wird das Einstellelement über Sprache angesteuert, die von einem Spracherkenner, insbesondere dem Spracherkennungssystem, ausgewertet wird. Die Eingangsgröße kann mittels Spracheingabe bestimmt werden.

Das Verfahren wird auch weitergebildet, indem eine vollständig automatisierte Bestimmung der Eingangsgröße in den folgenden Schritten durchgeführt wird:

Eine Leistungsfähigkeit des Rechners, auf dem das Spracherkennungssystem ablaufen soll, wird anhand eines Programms zur Leistungsermittlung bestimmt und als ein Leistungsindex abgespeichert. Unter Berücksichtigung des Leistungsindex werden die Systemparameter des Spracherkennungssystems automatisch eingestellt und dadurch wird eine leistungsfähige Spracherkennung unter z.B. Echtzeitbedingung gewährleistet.

Ein Programm zur Bestimmung der Leistungsfähigkeit des Rechners kann eine vorgegebene Schleife sein, die eine bestimmte Anzahl Iterationen durchläuft, wobei die Zeit für die Iterationen gemessen wird. Anhand einer Tabelle kann abhängig von der gemessenen Zeit der Leistungsindex bestimmt werden. Auch gibt es kommerziell oder als Freeware verfügbare



Programme, die die Leistungsfähigkeit des Rechners ermitteln und als eine Bewertungsgröße der Leistungsfähigkeit einen Leistungsindex ausgeben.

- 5 Auch wird eine Vorrichtung zur Spracherkennung angegeben, die ein Spracherkennungssystem aufweist und mit einem Mittel zur Einstellung einer Genauigkeit des Spracherkennungssystems ausgeführt ist, welches Mittel Systemparameter des Spracherkennungssystem aus einer Eingangsgröße umsetzt, also  
10 die Einstellung des Spracherkennungssystems und seiner zahlreichen Systemparameter anhand der Eingangsgröße vornimmt.

- Dabei ist es vorteilhaft, daß eine solche Einstellung der  
15 Systemparameter anhand der Eingangsgröße während des Betriebs des Spracherkennungssystems erfolgen kann. So wird für den Benutzer eine einfache Anpassung der Vielzahl von Systemparametern möglich.

- 20 Die Eingangsgröße ist in einer Weiterbildung automatisch bestimmbar. Dazu wird anhand einer Einrichtung zur Leistungsmessung des Rechners, auf dem das Spracherkennungssystem abläuft, ein sog. Leistungsindex ermittelt und zur Einstellung der Genauigkeit des  
25 Spracherkennungssystems benutzt.

- Eine andere Weiterbildung sieht vor, daß die Eingangsgröße durch ein Einstellelement vorgebar ist. Dazu sind eine Vielzahl möglicher Einstellelemente (Potentiometer, virtuelle  
30 Steuereinheiten auf dem Rechner, etc.) denkbar, deren Einstellung direkt die Genauigkeit des Spracherkennungssystems bestimmt.

- Vorteilhaft können Schritte des erfindungsgemäßen Verfahrens  
35 auf der angegebenen Vorrichtung durchgeführt werden.

Weiterbildungen der Erfindung ergeben sich auch aus den abhängigen Ansprüchen.

Anhand der folgenden Figuren werden Ausführungsbeispiele der Erfindung näher dargestellt.

Es zeigen

- Fig.1 eine Systemarchitektur für ein  
Spracherkennungssystem;
- Fig.2 ein Blockdiagramm mit Schritten eines Verfahrens zur  
Einstellung der Genauigkeit eines  
Spracherkennungssystems;
- Fig.3 ein Blockdiagramm, das eine Verknüpfung einer  
Eingangsgröße mit mindestens einem Systemparameter  
über eine Abbildungsvorschrift darstellt;
- Fig.4 ein Diagramm, das verschiedene mögliche  
Systemparameter des Spracherkennungssystems  
darstellt;
- Fig.5 eine Skizze, die den Einfluß der Systemparameter auf  
eine Zielgröße darstellt;
- Fig.6 ein Diagramm, das verschiedene Möglichkeiten zur  
Einstellung der Eingangsgröße darstellt;
- Fig.7 ein Spracherkennungssystem mit einem Mittel zur  
Einstellung der Genauigkeit.

In **Fig.1** ist allgemein eine Systemarchitektur für eine Spracherkennung (Spracherkennungssystem) dargestellt.

Voraussetzung für die Erkennung natürlich gesprochener Sprache ist ein geeigneter Formalismus zur

Wissensrepräsentation. Ein vollständiges  
Spracherkennungssystem umfaßt mehrere Verarbeitungsebenen.  
Dies sind insbesondere Akustik-Phonetik, Intonation, Syntax,  
Semantik und Pragmatik. In Fig.1 werden die  
5 Verarbeitungsebenen bei der Erkennung aufgezeigt.

Das natürliche Sprachsignal 101 gelangt in das  
Spracherkennungssystem. Dort wird in einer Komponente 102  
eine Merkmalsextraktion durchgeführt. Nach der  
10 Merkmalsextraktion werden anhand bekannter akustisch-  
phonetischer Einheiten 103 Sprachlaute erkannt (siehe Block  
104). Dabei handelt es sich um die Berechnung akustischer  
Distanzparameter. Nach der Sprachlauterkennung 104 erfolgt  
die lexikalische Decodierung (Worterkennung) in einem Block  
15 106 mit Hilfe des Aussprachemodells bzw. Wortlexikons 105 und  
daran anschließend eine Syntaxanalyse 108 mit Hilfe des  
Sprachmodells, das die Grammatik umfaßt, 107. Die  
Worterkennung 106 und die Syntaxanalyse 108 stellen die Suche  
nach einer Entsprechung für das Sprachsignal dar. Schließlich  
20 wird in einem Block 110 eine semantische Nachbearbeitung  
durchgeführt, wobei Kontextwissen und Pragmatik 109  
berücksichtigt werden und schließlich die vom  
Spracherkennungssystem erkannte Sprache 111 folgt.

25 In Fig.2 ist ein Blockdiagramm dargestellt, das Schritte  
eines Verfahrens zur Einstellung der Genauigkeit eines  
Spracherkennungssystems zeigt.

In einem Schritt 201 wird durch eine vorgebbare Eingangsgröße  
30 die Genauigkeit des Spracherkennungssystems bestimmt.  
Daraufhin wird in einem Schritt 202 anhand dieser  
Eingangsgröße für Systemparameter des Spracherkennungssystems  
Werte ermittelt. Schließlich wird in einem Schritt 203 das  
Spracherkennungssystem anhand der ermittelten Werte  
35 eingestellt.

**Fig.3** zeigt ein Blockdiagramm, das eine Verknüpfung einer Eingangsgröße über eine Abbildungsvorschrift mit mindestens einem Systemparameter darstellt.

Die erwähnte Eingangsgröße 301 wird anhand einer  
5 Abbildungsvorschrift 302 auf die Systemparameter SP 303 des  
Spracherkennungssystems abgebildet. Dabei wird vorzugsweise  
eine Eingangsgröße 301 mehreren Systemparametern über die  
Abbildungsvorschrift zugeordnet. Durch diese  
10 Abbildungsvorschrift 302 wird durch Vorgabe einer  
Eingangsgröße das Spracherkennungssystem angepaßt, also  
werden mehrere Systemparameter SP durch Veränderung einer  
Eingangsgröße 301 beeinflußt. Die Abbildungsvorschrift 302  
hat vorzugsweise die Form einer Tabelle, in der eine Spalte  
15 mögliche Eingangsgrößen 301 enthält, und in einer Zeile  
dieser Spalte der jeweiligen Eingangsgröße mehrere Werte für  
Systemparameter SP 303 zugeordnet werden. Die  
Abbildungsvorschrift 302 besteht im Suchen nach dem der  
Eingangsgröße 301 zugeordneten Eintrag (Zeile) in der Tabelle  
und in der Übergabe der gefundenen Werte für Systemparameter  
20 SP 303 an das Spracherkennungssystem.

**Fig.4** zeigt ein Diagramm, das verschiedene mögliche  
Systemparameter des Spracherkennungssystems darstellt. Die  
Systemparameter SP des Spracherkennungssystems, dargestellt  
25 in einem Block 401, umfassen mindestens einen der folgenden  
Parameter:

- a) Pruning-Schwelle 402;
- b) Histogramm-Pruning 403;
- c) akustische Vorausschau 404;
- 30 d) Vorausschaum im Sprachmodell 405;
- e) Schwelle für Distanzparameter 406.

Insgesamt sind weitere Systemparameter des  
Spracherkennungssystems zur Einstellung über die  
35 Eingangsgröße 301 denkbar, angedeutet durch den Block 407.

Fig.5 zeigt eine Skizze, die den Einfluß der Systemparameter auf eine Zielgröße darstellt.

- Die Systemparameter SP (siehe Block 501) nehmen Einfluß auf eine Zielgröße ZG (siehe Block 502). Wie oben beschrieben, gibt es mehrere Systemparameter SP1, SP2, usw., dargestellt in einem Block 503. Dabei nimmt jeder einzelne Systemparameter SPi (i=1,2,...) mittels eines für den jeweiligen Systemparameter SPi vorgesehenen Gewichts Gi (dargestellt in einem Block 504) Einfluß auf die Zielgröße ZG. Durch die Gewichtung der Systemparameter SP ist es möglich, je nach Systemparameter SP einen unterschiedlichen Einfluß auf die Zielgröße ZG zu nehmen.
- In Fig.6 ist ein Diagramm dargestellt, das verschiedene Möglichkeiten zur Einstellung der Eingangsgröße zeigt. Die Einstellung der Eingangsgröße, dargestellt in dem Block 601, erfolgt anhand von Bedienkomponenten des Rechners R (siehe Block 602), anhand von Regelungskomponenten 603 oder anhand des Rechners selbst (siehe Block 604). Der Rechner R umfaßt dazu mindestens ein Mittel zur Einstellung der Eingangsgröße, wie eine Tastatur 605, eine Maus 606, ein Touch-Pad 607 oder Spracheingabe 608 über das Spracherkennungssystem. Mögliche Komponenten zur Regelung, dargestellt in dem Block 603 sind ein Drehregler 609, ein Schieberegler 610 oder ein sonstiger Regler 611, vorzugsweise ein Potentiometer. Zusätzlich wird automatisch anhand eines Programms, das auf dem Rechner läuft, die Rechenleistung des Rechners ermittelt und die Eingangsgröße zur Einstellung des Spracherkennungssystems entsprechend bestimmt. Dadurch wird gewährleistet, daß ein automatisch eingestellter Rechner eine seiner Rechenleistung entsprechende Qualität bei der Spracherkennung gewährleistet. Automatisch wird ein Kompromiß gefunden zwischen hoher Qualität bei der Spracherkennung, unter Einbuße von Rechenleistung bzw. Echtzeiterkennung der Sprache, und schneller Spracherkennung mit entsprechend wenig Bedarf an

Rechenleistung allerdings deutlicher Qualitätseinbuße bei der Spracherkennung.

Fig.7 zeigt eine Vorrichtung aus einem Spracherkennungssystem 701 und einem Mittel zur Einstellung der Genauigkeit des Spracherkennungssystems 702.

Anhand des Mittels zur Einstellung der Genauigkeit 702 sind durch eine Eingangsgröße Systemparameter des Spracherkennungssystems bestimmt. Dazu wird die Eingangsgröße, vorzugsweise anhand einer vorgegebenen Tabelle, einer Vielzahl von Systemparametern (Pruning-Schwelle, Histogramm-Pruning, akustische Vorausschau, Vorausschau im Sprachmodell, Schwelle zur Auswahl zu berechnender Distanzparameter, etc.) zugeordnet.

Die Eingangsgröße ist wahlweise anhand eines Einstellelements 703 oder einer Einrichtung zur Leistungsbestimmung des Rechners 704 bestimmbar. Hierbei sei auch auf Fig.6 und die dort angegebenen Möglichkeiten zur Einstellung der Eingangsgröße verwiesen.

Im Rahmen dieses Dokuments wurden folgende Veröffentlichungen zitiert:

- [1] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 2, Seiten 13 bis 26.
- [2] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 3.3.3, Seiten 40 bis 43.
- [3] Volker Steinbiss, Bach-Hiep Tran, Hermann Ney: "Improvements in Beam Search. Proc. Intl. Conf. Speech and Language Processing, Yokohama 1994, Seiten 2143 bis 2146.
- [4] M. Niemöller, A. Hauenstein, E. Marschall, P. Witschel, U. Harke: "A PC-based Real-Time Large Vocabulary Continuous Speech Recognizer for German", Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; München 1997.
- [5] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 3.5.1, Seiten 65 bis 69.
- [6] S. Ortmanns, A. Eiden, H. Ney, N. Coenen: "Look-Ahead Techniques for Fast Beam Search", Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; München 1997, Seiten 1783 bis 1786.
- [7] E. Bocchieri: "Vector Quantization for the Efficient Computation of Continuous Density Likelihoods", Proc.

IEEE Intl. Conf. on Acoustics, Speech and Signal  
Processing; 1993, Seiten II-692 bis II-695.



Patentansprüche

1. Verfahren zur Spracherkennung,  
bei dem gesprochene Sprache anhand eines  
Spracherkennungssystems erkannt wird, wobei  
a) bei dem das Spracherkennungssystem auf einem Rechner  
abläuft;  
b) bei dem durch ein Programm zu einer  
Leistungsermittlung des Rechners ein Leistungsindex  
des Rechners bestimmt wird;  
c) bei dem die Eingangsgröße für das  
Spracherkennungssystem automatisch anhand des  
Leistungsindex bestimmt wird, wobei dabei eine  
Rechenleistung des Rechners automatisch auf die  
Genauigkeit des Spracherkennungssystems eingestellt  
wird.
2. Verfahren nach Anspruch 1,  
bei dem die Werte für die Systemparameter des  
Spracherkennungssystems bestimmt werden, indem gemäß  
einer Abbildungsvorschrift aus der Eingangsgröße die  
Werte ermittelt werden.
3. Verfahren nach Anspruch 2,  
bei dem die Abbildungsvorschrift anhand einer Tabelle  
umgesetzt wird.
4. Verfahren nach einem der vorhergehenden Ansprüche,  
bei dem die Einstellung während des Betriebs des  
Spracherkennungssystems durchgeführt wird.
5. Verfahren nach einem der vorhergehenden Ansprüche,  
bei dem das Spracherkennungssystem mindestens einen der  
folgenden Systemparameter umfaßt:  
a) Pruning-Schwelle;  
b) Histogramm-Pruning;  
c) akustische Vorausschau;

- d) Vorausschau im Sprachmodell;
- e) Schwelle zur Auswahl zu berechnender Distanzparameter.

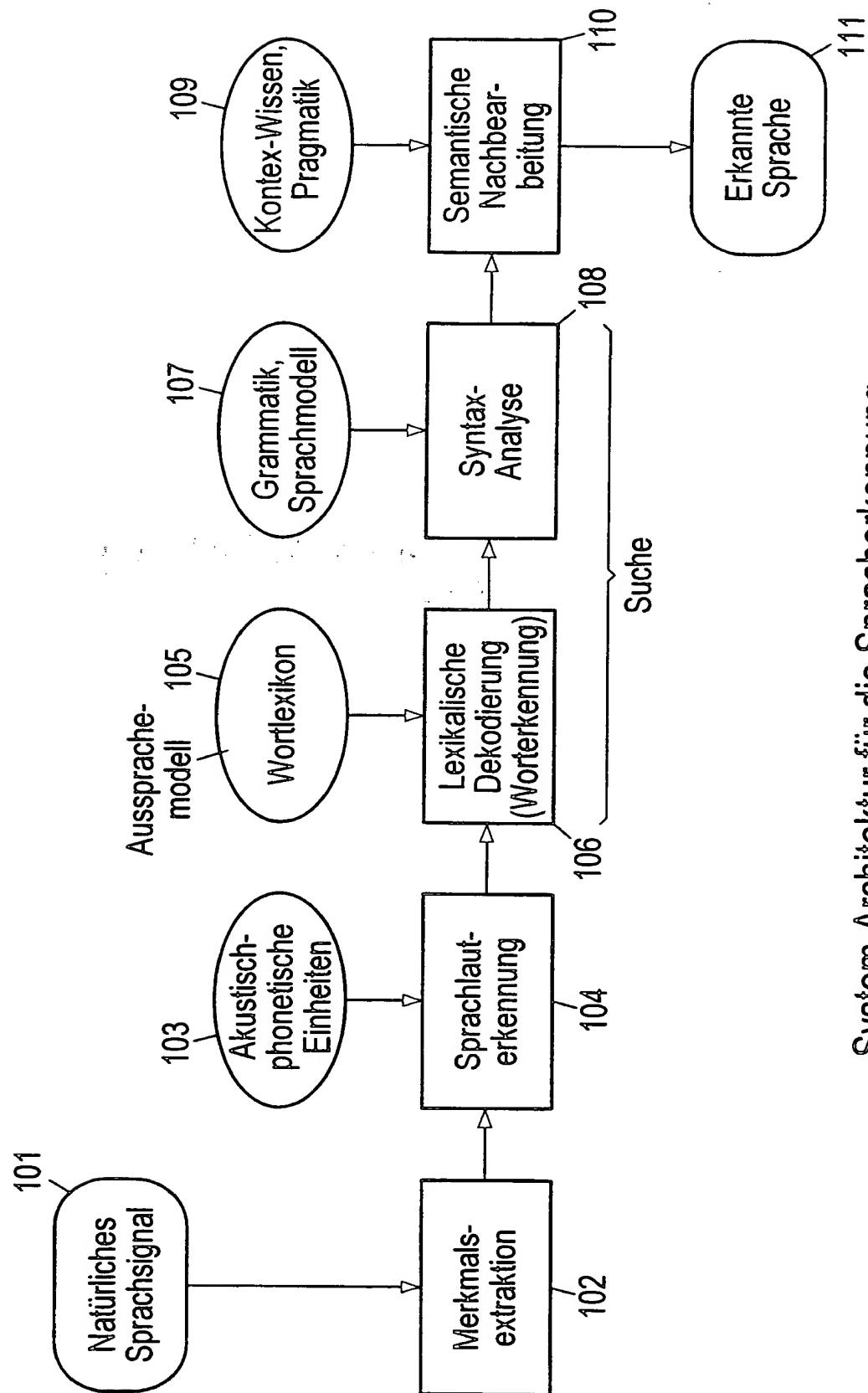
- 6. Verfahren nach Anspruch 5,  
5 bei dem mindestens einer der Systemparameter anhand der Eingangsgröße bestimmt wird.
- 7. Verfahren nach Anspruch 6,  
10 bei dem die Systemparameter gewichtet werden hinsichtlich ihres Einflusses auf jeweils eine Zielgröße.
- 8. Verfahren nach Anspruch 7,  
15 bei dem eine Zielgröße mindestens eine der folgenden Größen ist:
  - a) Genauigkeit des Spracherkennungssystems;
  - b) Geschwindigkeit des Spracherkennungssystems.
- 9. Verfahren nach Anspruch 7 oder 8,  
20 bei dem die Systemparameter gleich gewichtet werden.
- 10. Verfahren nach Anspruch 7 oder 8,  
bei dem die Systemparameter entsprechend einer vorgegebenen Gewichtungstabelle gewichtet werden.
- 25 11. Vorrichtung zur Spracherkennung,
  - a) bei der ein Spracherkennungssystem vorgesehen ist,
  - b) bei der ein Mittel zur Einstellung einer Genauigkeit des Spracherkennungssystems vorgesehen ist, das derart
- 30 eingerichtet ist, daß Systemparameter des Spracherkennungssystems einstellbar sind, wobei die Systemparameter anhand einer Eingangsgröße ermittelbar sind.
- 35 12. Vorrichtung nach Anspruch 11,  
mit einer Einrichtung zur Leistungsmessung, die derart

eingerrichtet ist, daß die EingangsgroÙe automatisch  
bestimmbar ist.

**This Page Blank (uspto)**

1/4

FIG 1



System-Architektur für die Spracherkennung

**This Page Blank (uspto)**

2/4

FIG 2

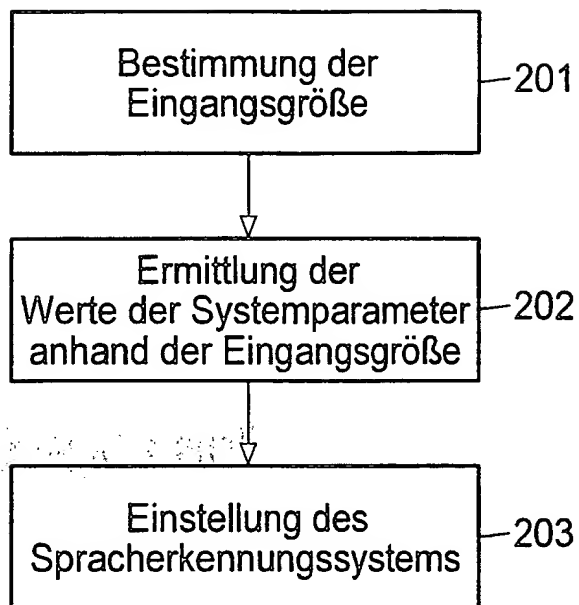
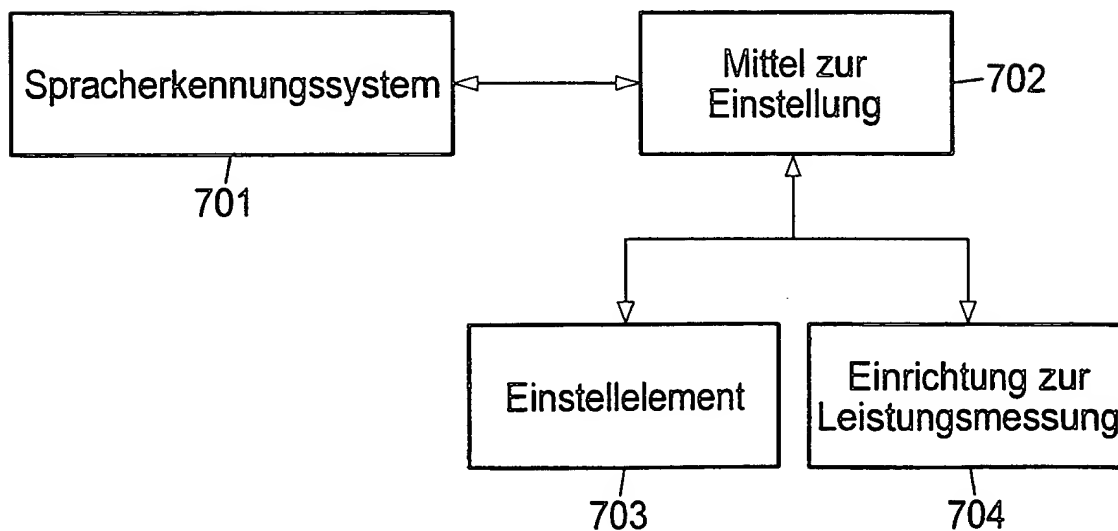


FIG 7



**This Page Blank (uspto)**



3/4

FIG 3

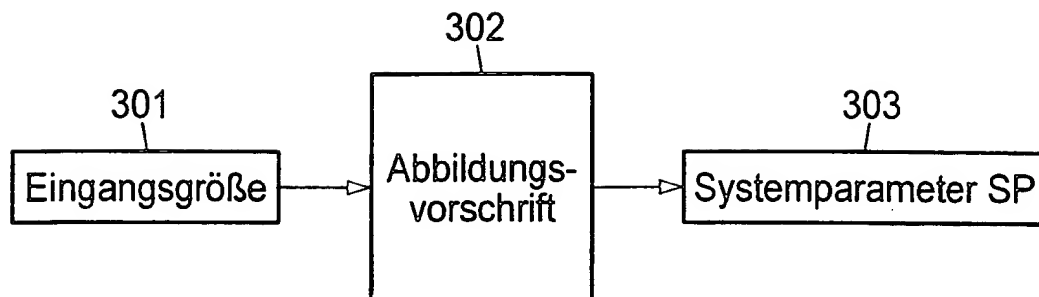
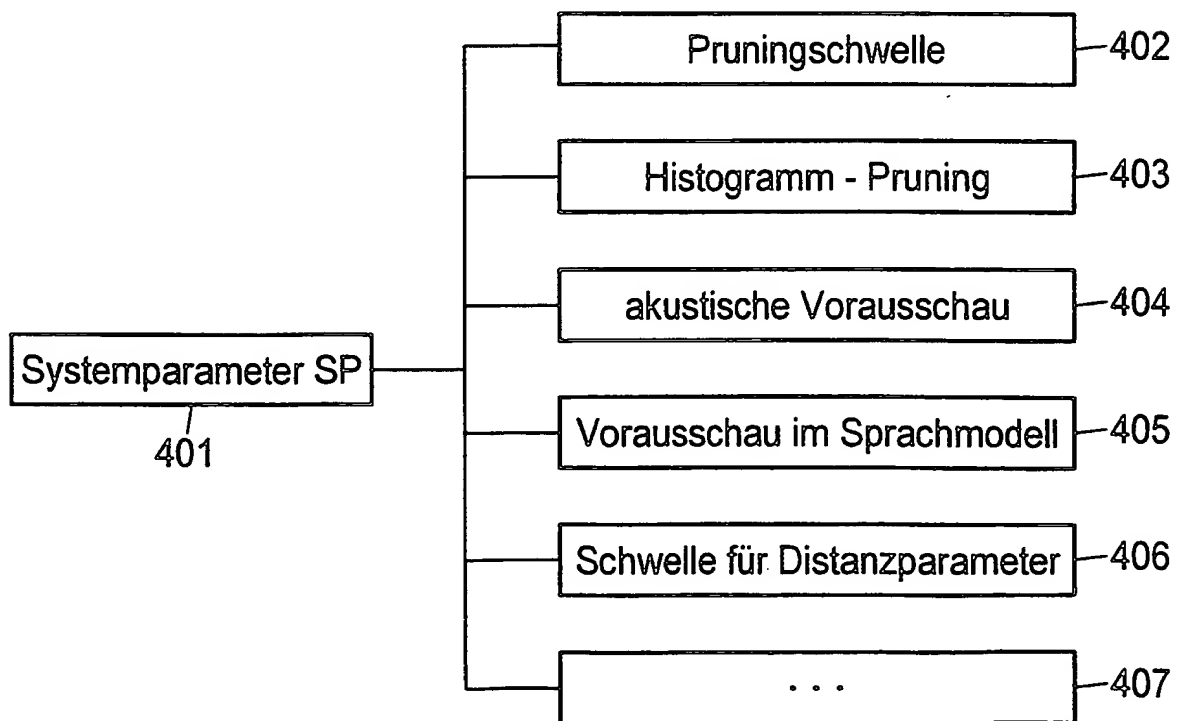


FIG 4



This Page Blank (uspto)

4/4

FIG 5

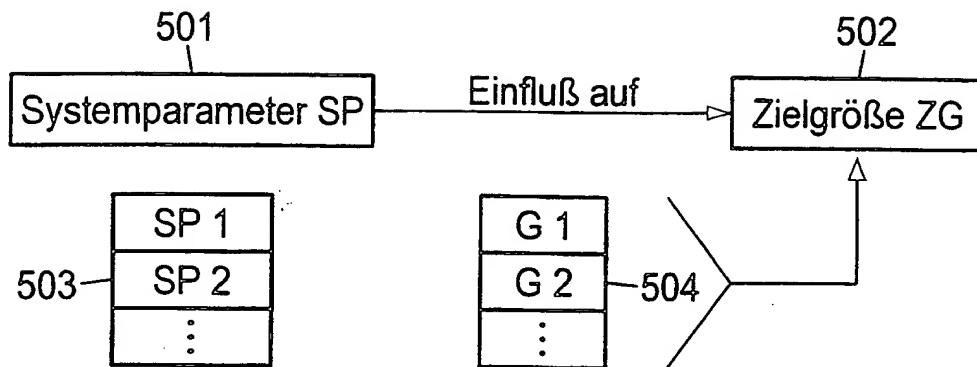
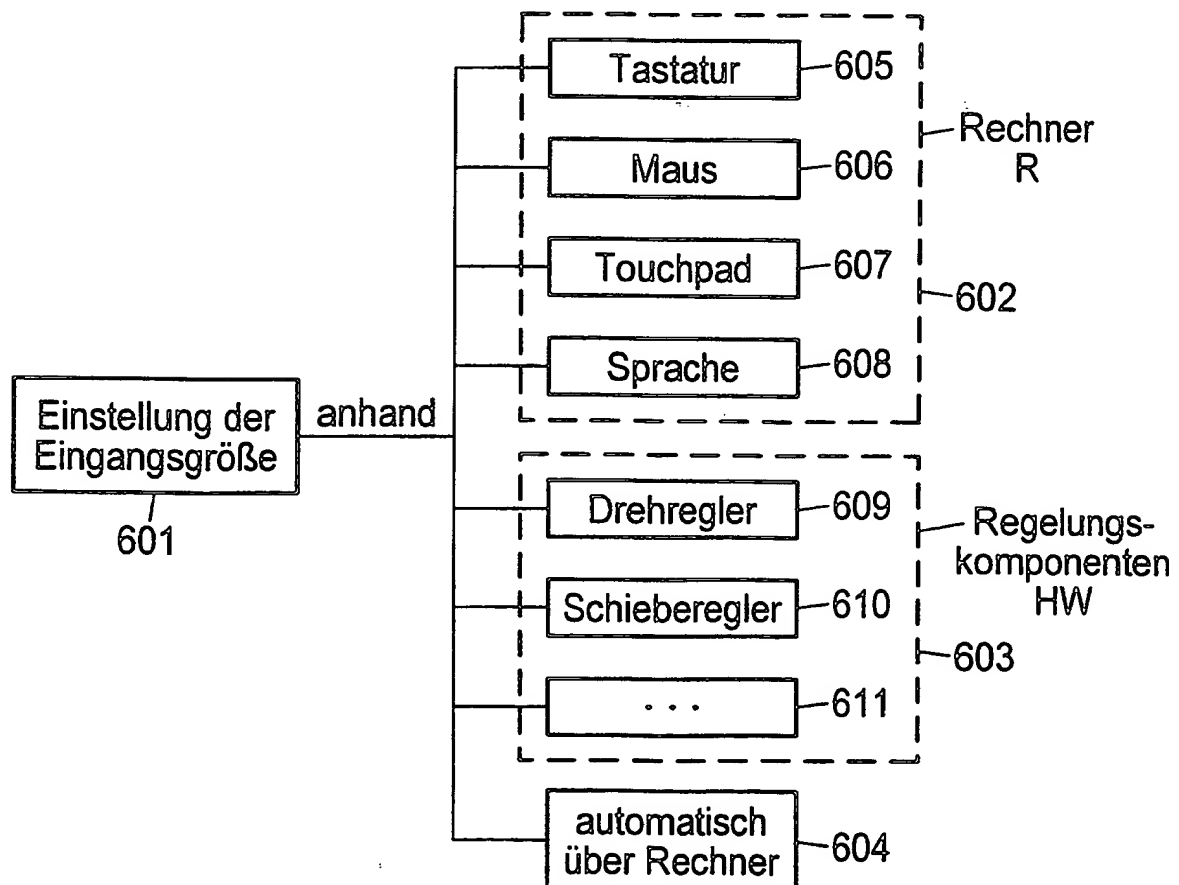


FIG 6



**This Page Blank (uspto)**

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/DE 98/03366

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 G10L5/06 G10L3/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 682 464 A (SEJNOHA VLADIMIR) 28 October 1997 see abstract; claims 7,9 ---	1,11,12
P,X	WO 98 22936 A (NETIX INC T) 28 May 1998 see page 30, line 7 - line 21; claims 8-12	1,11
Y	---	5
Y	BOCCHIERI E: "VECTOR QUANTIZATION FOR THE EFFICIENT COMPUTATION OF CONTINUOUS DENSITY LIKELIHOODS" SPEECH PROCESSING, MINNEAPOLIS, APR. 27 - 30, 1993, vol. 2, 27 April 1993, pages II-692-695, XP000427884 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS see paragraph 5 --- -/--	5

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

13 April 1999

Date of mailing of the international search report

20/04/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Van Doremalen, J

# INTERNATIONAL SEARCH REPORT

Intern. Application No

PCT/DE 98/03366

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>ORTMANNS S ET AL: "LOOK-AHEAD TECHNIQUES FOR FAST BEAM SEARCH"</p> <p>1997 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, SPEECH PROCESSING, DIGITAL SIGNAL PROCESSING MUNICH, APR. 21 - 24, 1997, vol. 3, 21 April 1997, pages 1783-1786, XP000735006</p> <p>INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS</p> <p>see abstract</p> <p>-----</p>	5

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/DE 98/03366

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5682464 A	28-10-1997	US 5386492 A	31-01-1995
		US 5546499 A	13-08-1996
		EP 0648366 A	19-04-1995
		JP 7508360 T	14-09-1995
		WO 9400836 A	06-01-1994
		WO 9533259 A	07-12-1995
WO 9822936 A	28-05-1998	AU 5359498 A	10-06-1998

**This Page Blank (uspto)**



A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
IPK 6 G10L5/06 G10L3/00

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 6 G10L

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie <sup>a</sup>	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	US 5 682 464 A (SEJNOHA VLADIMIR) 28. Oktober 1997 siehe Zusammenfassung; Ansprüche 7,9 ---	1, 11, 12
P, X	WO 98 22936 A (NETIX INC T) 28. Mai 1998 siehe Seite 30, Zeile 7 - Zeile 21; Ansprüche 8-12	1, 11
Y	---	5
	---	
	---/---	



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

<sup>a</sup> Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderscher Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderscher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

13. April 1999

Absendedatum des internationalen Recherchenberichts

20/04/1999

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Van Doremalen, J

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	BOCCHIERI E: "VECTOR QUANTIZATION FOR THE EFFICIENT COMPUTATION OF CONTINUOUS DENSITY LIKELIHOODS" SPEECH PROCESSING, MINNEAPOLIS, APR. 27 - 30, 1993, Bd. 2, 27. April 1993, Seiten II-692-695, XP000427884 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS siehe Absatz 5 ----	5
Y	ORTMANN S ET AL: "LOOK-AHEAD TECHNIQUES FOR FAST BEAM SEARCH" 1997 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, SPEECH PROCESSING, DIGITAL SIGNAL PROCESSING MUNICH, APR. 21 - 24, 1997, Bd. 3, 21. April 1997, Seiten 1783-1786, XP000735006 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS siehe Zusammenfassung -----	5